AWARD NUMBER: W81XWH-14-1-0110

TITLE: A Molecular Framework for Understanding DCIS

PRINCIPAL INVESTIGATOR: Gregory Hannon

CONTRACTING ORGANIZATION: The Chancellor, Masters and Scholars of the University of Cambridge, Clara East, The Old Schools, Cambridge CB2 1TN

REPORT DATE: October 2016

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**
October 2016	Annual	9/30/15–10/29/16

4. TITLE AND SUBTITLE

A Molecular Framework for Understanding DCIS

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-14-1-0110

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Gregory Hannon PhD/ Clare Rebbeck

E-Mail: greg.hannon@cruk.cam.ac.uk

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

CRUK Cambridge Institute
Li Ka Shing Centre
Robinson Way
Cambridge
CB2 ORE

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This project centers on creating a molecular framework of DCIS (ductal carcinoma in situ). DCIS is considered to be the precursor to Invasive Ductal Carcinoma (IDC), the most common form of breast cancer. IDC accounts for 80% of all breast cancers, predominantly affecting women aged 55 and older; however, at least a third of women with IDC are diagnosed before they reach 55. Not all patients with DCIS will develop IDC however, we are looking for ways to better predict those patients that need life-saving treatment, and separate these from those patients who are less at risk.

So far we have over 3000 lesions dissected from 70 freshly frozen patient biopsies, these have been annotated by our pathologist and prepared to be taken on for sequencing. The tissue includes DCIS, IDC, stroma adjacent to DCIS/IDC and normal tissue. We have initiated the RNA sequencing from these samples and also the DNA sequencing

15. SUBJECT TERMS
DCIS, IDC, LCM, RNAseq, DNAseq, evolution

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	7	**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

A Molecular Framework for understanding DCIS

Award No. W81XWH-14-1-0110
Annual report year 2

1
Introduction

This project centers on creating a molecular framework of DCIS (ductal carcinoma in situ). DCIS is considered to be the precursor to Invasive Ductal Carcinoma (IDC), the most common form of breast cancer. IDC accounts for 80% of all breast cancers, predominantly affecting women aged 55 and older; however, at least a third of women with IDC are diagnosed before they reach 55.

Utilizing a unique bank of frozen mammary biopsies, containing samples with DCIS alone, and a combination of DCIS and IDC, we aim to profile both DCIS and related tissue components. It is our aim to sample the ~300 biopsies, and compare both by RNA seq, and whole genome amplification, DCIS lesions, within, and between patients, and see how these may be correlated with IDC lesions. We also intend to look for changes in the stroma between those patients that present with IDC and those that do not. This work aims to identify characteristics that may be suggestive of a patients' likelihood of progressing from DCIS to IDC, with the purpose of reducing the need for over treatment for this disease.

2
Keywords

Ductal carcinoma in situ, DCIS, Invasive ductal carcinoma, IDC, RNA, DNA, Copy number, Laser capture microscope, LCM

3.
Accomplishments

Aim 1. The evolution of DCIS.

Task 1. Sample collection and annotation
Task 2. Sample choice from frozen bank.

We have received 80 samples from the frozen bank now and have processed 70 of these so far. These include pure DCIS and also mixed DCIS and IDC samples. We selected samples that had 5 or more DCIS legions for this Aim as these will be more informative for looking at the evolution of DCIS.

Task 3. Laser capture of frozen samples for characterization

From each of the 70 samples we have dissected material for DNA, however we have material from 18 patients for characterization (based on having 5 or more DCIS legions). We have selected DCIS legions, IDC regions, normal epithelium, where present, Atypical epithelium, Solid DCIS, papillary DCIS, benign epithelium, and stroma (as far away from DCIS or IDC regions as possible). The table below represents the distribution across patients, with a total of 214 legions including the normal and variants of epithelium. .

Number of DCIS legions	Number of samples	Number of samples with IDC	Number of IDC legions per sample
5	4	2	4,4
6	7	3	8,2,4
7	6	3	5,7,5
8	1	0	
9	1	1	5
10	1	1	6
11	0	0	
12	0	0	
13	1	0	
Total	**144**	**10**	**50**

Task 4. Exome capture and sequencing

We initiated work on this using the Nextera Exome Capture kit, however on the couple of samples we used, this did not prove successful, as there was a very low distribution of probes represented. Having investigated the costs and what is needed to get deep enough coverage for accurately calling CNVs and SNVs, we decided to make use of the X10 sequencing machine at the NYGC and do whole genome sequencing instead. A trial run with this demonstrated that the Whole Genome Sequencing kit that we were using (and other kits on the market) was only compatible with sequencing machines after the DNA had been sheared (resulting in removal of end primers). This was not efficient with sequencing on the X10, as reads are generally longer and shearing would result in very short reads. We therefore established a new protocol based on old school molecular biology, where by we literally chewed the primers off the ends of DNA strands after amplification with the WGA kit, this then allowed us to attach the primers for sequencing (this was somehow hindered without removal of the WGA primers). This pipeline proved very effective and all 214 samples have now just finished being sequenced on the X10.

Task 5. Analyze Exome capture data

Data from the X10 sequencing is still being processed however concordance analysis has been carried out on over half of the samples. This looks for any discrepancies between a “normal” sample and its paired “tumor” sample. Pairs generally have over 90% concordance, however this analysis has proved useful as it identified a misread tube label and thus allows us to correct such errors.

Initial analysis on CNV data has been carried out on 4 patients thus far and shows that there are differences to be seen between DCIS legions within the same patient. Further, more indepth analysis will be carried out on the phylogeny of the DCIS and IDC legions and if there are any associations between the differences we see in the DNA data and the differences we see in the RNA data. This is being carried out together with the bioinformaticians at the NYGC and we will seek further analysis from groups here at Cambridge who specialize in tumor evolution.

Aim 2. <u>A transcriptional landscape of early breast cancer.</u>

Task 6. Sample choice from frozen bank.
- choose samples for pure DCIS and DCIS with microinvasion/IDC

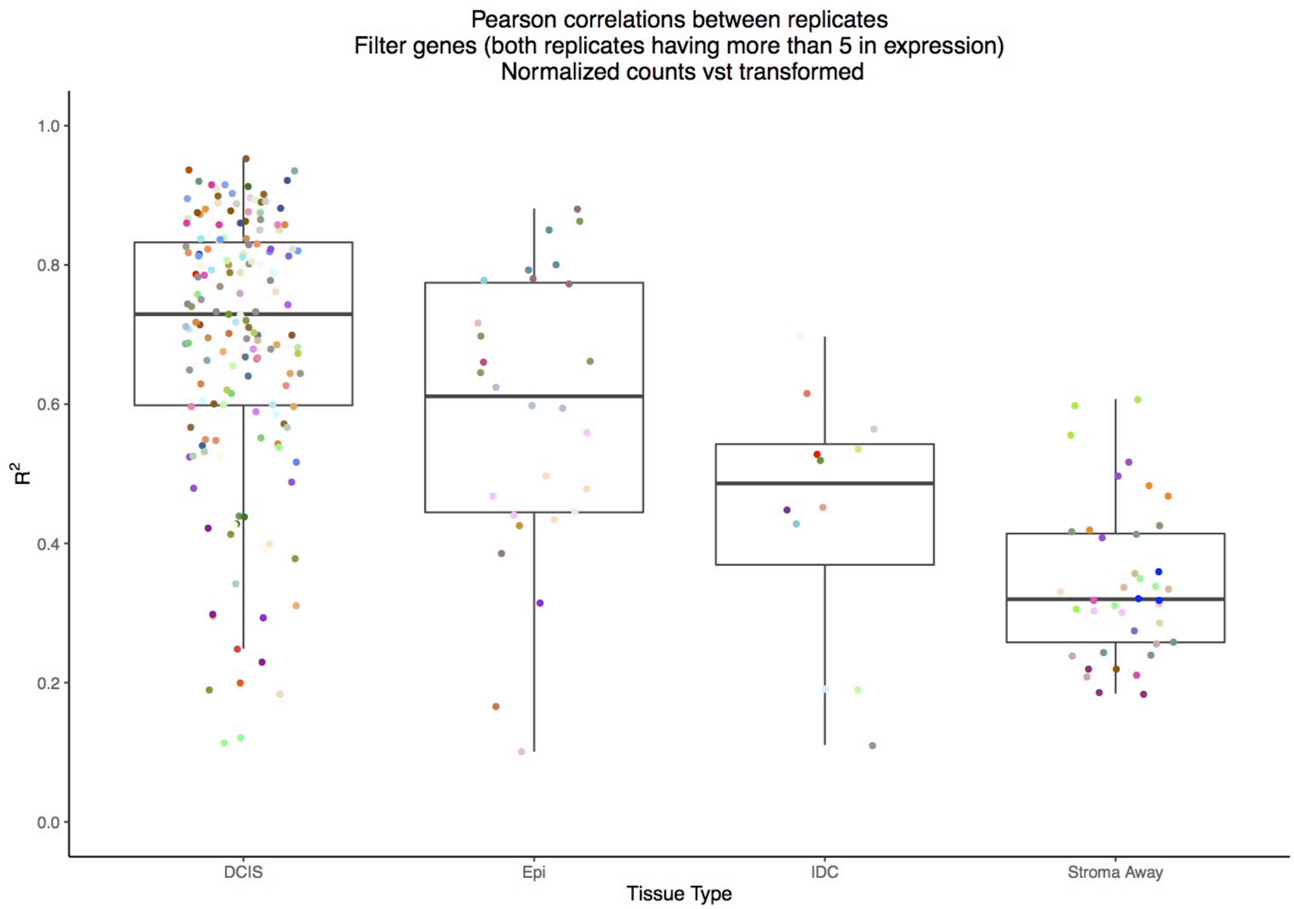
Task 7. Laser capture of frozen samples for characterization
We have received 80 samples from the frozen bank now and have processed 70 of these so far. For each sample the following regions are annotated by Joe (the pathologist) and dissected in triplicate for RNA: DCIS, IDC, normal epithelium, Atypical epithelium, Solid DCIS, papillary DCIS, benign epithelium, areas of high immune infiltration, stroma adjacent to DCIS, stroma adjacent to IDC and stroma away (as far away from DCIS or IDC regions as possible). This has provided over 3000 legions. This Task is still on going.

Task 8. RNAseq library construction

Approximately 800 RNA seq libraries have been sequenced. Currently we have been focusing on DCIS and IDC and normal and other epithelium and are prioritizing these for sequencing now. This task is still on going.

Task 9. Analyze RNAseq datasets

Thus far we have analyzed 318 DCIS, 95 IDC, 69 benign/normal epithelium and 91 stroma away libraries. For quality control samples with a Gene Assignment of < 15% with % of Uniquely mapped reads < 20, are removed from the group analysis. This has resulted in 122 libraries being removed. 1.5% DCIS, 2.1% IDC, 2.6% epithelium and 3.8% stroma away.



Opportunities for training and professional development

Nothing to report (not intended for training)

Results disseminated to communities of interest

Nothing to report

4.
Impact

Nothing to report

5.
Changes / problems

Nothing to report

6.
Products

Nothing to report

7.
Participants & other collaborating organizations

Individuals worked on the project

Name : Greg Hannon
Project Role: Initiating PI – contributed to project design and liaising with bioinformatics team
Nearest person month worked: 1 CM (10% x 13 months)
Funding support: CR-UK and Royal society

Name : Clare Rebbeck
Project Role: Co-PI – contributed to project design, staining strategy, dissecting with the LCM, RNA and DNA library preparation and liaising with Bioinformatics team and pathologist.
Nearest person month worked: 13 CM

Name : Jian Xian
Project Role: senior research assistant - contributed to dissecting with the LCM and RNA and DNA library preparation
Nearest person month worked: 12 CM
Funding support: CR-U

Name : Laurence de Torrente
Project Role: Bio informaticitian - contributed to data processing and data analysis
Nearest person month worked: 12 CM

Name : Martin Fabry
Project Role: student – contributed to DNA library preparation
Nearest person month worked: 12 CM

Name : Sophie Watcham
Project Role: research technician - contributed to DNA library preparation
Nearest person month worked: 3 CM

Change in active support since last report

Nothing to report (this is the first reporting period)

Other organizations involved as partners

Duke university – collaboration to provide tissue samples and clinical annotation; as detailed in the grant application.

New York Genome Center – Collaboration with the bioinformatics team to analysis the data; As detailed in the grant application